



調査報告書 | 2025年7月

生成AIツールの種類と 動向調査報告

本報告書では、最新の生成AIツールの分類、主要カテゴリ、技術進歩、応用事例、そして将来展望について体系的に解説します。

¹/₂ 目次

1. タイトル
2. 目次
3. 生成AIの概要と定義
4. 歴史的背景と現在の重要性
5. 生成AIツールの分類体系
6. テキスト生成ツール（LLM）
7. 画像・動画生成ツール
8. コード生成・音声生成・マルチモーダルツール
9. 主要ツールの比較分析
10. 横断的応用事例
11. 新興トレンドと課題
12. 結論と将来展望



💡 生成AIの概要と定義

生成AIとは

生成AIは、既存データを単に分析・分類するだけでなく、テキスト、画像、音声、動画、コードなどの**新しいかつ現実的なコンテンツを生成**できる人工知能モデルです。

識別AIが予測やカテゴリ分類に重点を置くのに対し、生成AIは学習データに似た新しいインスタンスの**創造的な生成**に焦点を当てています。

生成AIと識別AIの違い

- 識別AI：データを分析し、分類や予測を行う（例：画像認識、スパム検出）
- 生成AI：
学習したパターンを基に新しいコンテンツを創造する（例：文章作成、画像生成）



創造性



パターン学習



新規性

生成AIツールの種類と動向調査報告

🕒 歴史的背景と現在の重要性

生成モデルの発展

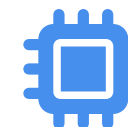
- **2014年**：敵対的生成ネットワーク（GANs）の登場 - 生成器と識別器の競合により高品質な生成を実現
- **2017年**：Transformerアーキテクチャの開発 - 自己注意機構により大規模言語モデル（LLM）の基盤に
- **2020年以降**：拡散モデルの台頭 - 高品質で多様な画像・動画生成を可能にし、GANsを凌駕

現在の重要性

- 急速な市場拡大：多額の投資と企業導入の加速
- 創造から最適化へ：AIが解決できる問題の種類の根本的变化
- 汎用技術（GPT）：電気やインターネットのように経済全体に水平的影響
- アクセシビリティ：API化とオープンソース化による広範な利用機会



急成長市場



計算能力の向上



技術の相乗効果

生成AIツールの分類体系

モダリティによる分類

生成AIツールは、生成するデータの**種類（モダリティ）**に基づいて分類できます：

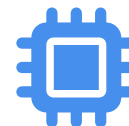
- テキスト：文章、記事、対話、翻訳など
- 画像/動画：デジタルアート、グラフィック、メディア制作
- コード：ソフトウェア開発、プログラミング支援
- 音声：音声合成、音楽作曲、効果音生成



テキスト



画像・動画



マルチモーダル

機能による分類と最新トレンド

各モダリティ内で、ツールは特定の**機能**に特化しています：

- コンテンツ作成：新規コンテンツの生成
- 要約/変換：既存コンテンツの処理と変換
- デザイン：視覚的コンテンツのスタイル付け

最新トレンド：境界が曖昧になり、**マルチモーダル**ツールが複数の種類のデータを同時に処理・生成できるようになっています

🗨️ テキスト生成ツール（LLM）

大規模言語モデル（LLM）とは

LLMはTransformerアーキテクチャを基盤とし、**人間のようなテキストを理解・生成**する能力を持つAIモデルです。膨大なデータで事前学習され、文脈を理解し、一貫性のある自然な文章を生成します。

主要なLLMとその特徴

ChatGPT（OpenAI）

GPT-3.5/4をベースとした汎用会話モデル。コンテンツ作成、質問応答、コード生成など幅広い能力を持つ。

Gemini（Google）

マルチモーダル能力に優れ、特にGemini 1.5 Proは100万トークンを超える長文コンテキスト処理が可能。

Claude（Anthropic）

安全性と倫理面に重点を置き、長文処理に強み。企業向けアプリケーションに適している。

Llama（Meta）

オープンソースモデルとして研究やカスタム開発に活用。様々なサイズと言語に対応。



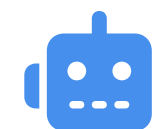
コンテンツ作成

記事、マーケティング素材、創作



翻訳・要約

多言語対応、文書要約



会話型AI

チャットボット、カスタマーサービス



分析・調査

情報抽出、研究支援

画像・動画生成ツール

拡散モデルの台頭

現代の画像・動画生成ツールの多くは**拡散モデル（Diffusion Models）**を基盤としています。これらは、ノイズから段階的に画像を生成する手法で、GANsに比べて高品質かつ多様性のある出力を実現します。

テキストプロンプトから精密に制御された視覚コンテンツを生成する能力は、クリエイティブワークフローを根本から変革しています。

主要な画像・動画生成ツール

- DALL-E（OpenAI）：高精度なテキスト記述からの画像生成、芸術的なスタイル制御、編集機能
- Midjourney：芸術性に優れた画像生成、独自の美的スタイルと細部表現の優位性
- Stable Diffusion（Stability AI）：オープンソース、カスタマイズ性、ローカル実行の可能性
- RunwayML：テキストからの動画生成、映像編集機能、創造的な映像制作ツール



デジタルアート



動画制作



デザイン

✂️ コード生成・音声生成・マルチモーダルツール

コード生成ツール

GitHub CopilotやAmazon CodeWhispererなどのツールは、大規模なコードリポジトリで学習したLLMを活用し、以下の機能を提供します：

- 自然言語からの自動コード生成
- インテリジェントなコード補完・サジェスト
- バグ検出とデバッグ支援



コード生成

音声生成ツール

ElevenLabsなどのツールは、超リアルな音声合成技術で以下を実現します：

- 自然な抑揚と感情表現を持つTTS（テキスト読み上げ）
- 高精度な音声クローニング（特定話者の声を再現）
- 多言語対応と声質カスタマイズ



音声合成

マルチモーダルツール

Google GeminiやOpenAI GPT-4Vなどのツールは複数の情報形式を処理・生成します：

- テキストと画像の統合理解と生成
- 長文コンテキストウィンドウによる動画処理（Gemini 1.5 Pro）
- クロスモーダル検索と分析（異なる形式間の変換）



マルチモーダル

主要ツールの比較分析

主要な生成AIツールの機能、用途、基盤技術の比較

ツール名	開発元	主要モダリティ	主要機能	主要用途	基盤モデル
 ChatGPT	OpenAI	テキスト生成	高度な会話、コンテンツ作成、要約、Q&A	コンテンツ作成、顧客サービス、教育	LLM (GPTシリーズ)
 Midjourney	Midjourney Inc.	テキスト-画像	高品質画像生成、芸術的スタイル	デジタルアート、コンセプトアート	拡散モデル
 Gemini 1.5 Pro	Google	マルチモーダル	長文コンテキスト、クロスモーダル理解	企業ソリューション、複雑データ分析	マルチモーダル Transformer
 GitHub Copilot	Microsoft/GitHub	コード生成	コード補完、自然言語からコード生成	ソフトウェア開発、プログラミング教育	LLM (Codex)
 ElevenLabs	ElevenLabs	音声生成	リアルな音声合成、音声クローン	オーディオブック、ポッドキャスト	Transformer ベース
 Stable Diffusion	Stability AI	テキスト-画像	オープンソース、カスタマイズ性	デジタルアート、デザイン、研究	拡散モデル
 Claude	Anthropic	テキスト生成	長文処理、安全性と倫理重視	企業向け応用、コンテンツ生成	LLM

🛠️ 横断的応用事例

業界横断的な生成AIの活用

生成AIは様々な産業で広範な応用が進み、タスク自動化だけでなく**人間の創造性と生産性を拡張**する役割を担っています。

✍️ クリエイティブ産業

- デジタルアートとイラストレーション制作
- マーケティングコピーや広告コンテンツの生成
- 音楽作曲とオーディオコンテンツ制作

💓 ヘルスケア

- 創薬プロセスの加速化
- 診断支援と医療画像の分析
- 個別化医療と治療法設計

💻 ソフトウェア開発

- コード自動生成とバグ修正支援
- 開発プロセスの高速化とプロトタイピング
- テスト自動化とドキュメント作成

📈 金融・ビジネス

- 市場分析とリスク評価
- 顧客サービスの自動化
- 不正検出と予防



教育

個別学習コンテンツ
言語学習支援



法律

法務調査自動化
文書レビュー



セキュリティ

脆弱性検出
防御シミュレーション

生成AIツール

 Genspark で作成

📌 新興トレンドと課題

新興技術トレンド

AIエージェント

ツールや環境と自律的に対話し、複雑な多段階タスクを実行できるAIシステム。単一クエリから持続的で目標指向の処理へと進化しています。

エッジAI

生成AIモデルをデバイス上で直接実行することで、リアルタイム処理、低遅延、プライバシー強化を実現。計算をクラウドからデバイスへと分散化します。

オープンソースAI

MetaのLlamaシリーズなど、強力なオープンソースモデルの登場。AIへのアクセスを民主化し、イノベーションを促進しています。

課題と懸念事項

幻覚と事実確認

生成AIの「幻覚」問題。正確でない、または無意味な情報を生成する課題と、RAGなどの解決アプローチ。

倫理的課題と偏見

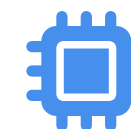
偏ったデータで訓練されたモデルによる社会的偏見の永続化や増幅。公平性の確保が重要課題です。

知的財産と著作権

著作権保護データによる訓練やIP権を侵害する可能性のあるコンテンツ生成に関する複雑な法的問題。



自律性



分散化



オープン化



ガバナンス

📌 結論と将来展望

現状の総括

生成AIツールはテキスト、画像、コード、音声、マルチモーダルといった多様なモダリティにわたって急速に進化しています。TransformerやLLM、拡散モデルなどの基盤技術が、このイノベーション波を支えています。

ほぼすべての産業において、生成AIがワークフローを変革し、新しい機会を創出していることが明らかになっています。

将来の方向性

- 専門化とカスタマイズ：カスタムGPTや長文コンテキスト機能に見られる特化型AI
- 倫理と責任：偏見、知的財産権、幻覚などの問題への継続的な対応
- 人間とAIの協業：AIを完全な代替ではなく、人間の創造性と生産性を拡張するパートナーとして活用
- 自律・統合・アクセス性：AIエージェント、エッジAI、オープンソースモデルの継続的な発展



人間とAIの協業



技術革新



倫理と責任